

Conditional negative association for competing urns*

J. Kahn and M. Neiman

Rutgers University and UCLA

email: jkahn@math.rutgers.edu; neiman@math.ucla.edu

Abstract

We prove conditional negative association for random variables $\mathbf{x}_j = \mathbf{1}_{\{|\sigma^{-1}(j)| \geq t_j\}}$ ($j \in [n] := \{1, \dots, n\}$), where $\sigma(1), \dots, \sigma(m)$ are i.i.d. from $[n]$. (The $\sigma(i)$'s are thought of as the locations of balls dropped independently into urns $1, \dots, n$ according to some common distribution, so that, for some threshold t_j , \mathbf{x}_j is the indicator of the event that at least t_j balls land in urn j .) We mostly deal with the more general situation in which the $\sigma(i)$'s need not be identically distributed, proving results which imply conditional negative association in the i.i.d. case. Some of the results—particularly Lemma 8 on graph orientations—are thought to be of independent interest.

We also give a counterexample to a negative correlation conjecture of D. Welsh, a strong version of a (still open) conjecture of G. Farr.

1 Introduction

Competing urns refers to the experiment in which m balls are dropped, randomly and independently, into urns $1, \dots, n$. Formally, we have a random $\sigma : [m] \rightarrow [n]$ (where $[m] = \{1, \dots, m\}$) with the $\sigma(i)$'s independent. We then take \mathbf{x}_j to be the indicator for occupation of urn j and are interested in the law, μ , of $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ (a measure on $\{0, 1\}^n$). In the traditional case where the balls are identical (i.e. the $\sigma(i)$'s are i.i.d.) we call μ an *urn measure*, or, for emphasis, an *ordinary urn measure*. More generally, setting $B_j = |\sigma^{-1}(j)|$, we may consider thresholds t_1, \dots, t_n , and let \mathbf{x}_j be the indicator of $\{B_j \geq t_j\}$; for i.i.d. balls, we then call the law of $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ a *threshold urn measure*. When the balls are not required to be identical we speak of *generalized urn measures* and *generalized threshold urn measures*.

We are interested in correlation properties of these measures, but before proceeding further need to briefly recall a few definitions. A fuller version of the following discussion is given in [10], and further background and motivation may be found e.g. in [15].

Recall that events \mathcal{A}, \mathcal{B} in a probability space are *positively correlated*—we write $\mathcal{A} \uparrow \mathcal{B}$ —if $\Pr(\mathcal{A}\mathcal{B}) \geq \Pr(\mathcal{A})\Pr(\mathcal{B})$, and *negatively correlated* ($\mathcal{A} \downarrow \mathcal{B}$) if the reverse inequality holds.

We will be interested in measures on finite product spaces $\Omega = \prod_{i=1}^n \Omega_i$ with each Ω_i a chain (totally ordered set), often simply $\{0, 1\}$. We use $\mathcal{M}(\Omega)$, or simply \mathcal{M} , for the set of probability measures on Ω , and \mathcal{M}_S for $\mathcal{M}(\{0, 1\}^S)$. We will occasionally identify $\{0, 1\}^S$ with 2^S (subsets of S) ordered by inclusion in the usual way.

Recall that an event $\mathcal{A} \subseteq \Omega$ is *increasing* (really, nondecreasing) if $x \geq y \in \mathcal{A} \Rightarrow x \in \mathcal{A}$ (where Ω is endowed with the product order), and similarly for *decreasing*. For real-valued random

AMS 2000 subject classification: 60C05, 05A20

Key words and phrases: competing urns, correlation inequalities, conditional negative association, log-concavity

* Supported by NSF grant DMS0701175.

variables X, Y , write $X \downarrow Y$ if

$$\{X \geq s\} \downarrow \{Y \geq t\} \quad \forall s, t \in \mathfrak{R}, \quad (1)$$

or, equivalently, if

$$\mathbb{E}f(X)g(Y) \leq \mathbb{E}f(X)\mathbb{E}g(Y) \text{ for all increasing } f, g : \mathfrak{R} \rightarrow \mathfrak{R}. \quad (2)$$

(N.B. this differs from the usage in [15]. Of course $X \uparrow Y$ means the reverse inequalities hold, but we don't need this.)

Say $i \in [n]$ *affects* $\mathcal{A} \subseteq \Omega$ if there are $\eta \in \mathcal{A}$ and $\tau \in \Omega \setminus \mathcal{A}$ with $\eta_j = \tau_j \forall j \neq i$, and write $\mathcal{A} \perp \mathcal{B}$ if no coordinate affects both \mathcal{A} and \mathcal{B} . Then $\mu \in \mathcal{M}$ is *negatively associated* (or has *negative association*; we use “NA” for either) if $\mathcal{A} \downarrow \mathcal{B}$ whenever \mathcal{A}, \mathcal{B} are increasing and $\mathcal{A} \perp \mathcal{B}$. We say μ has *negative correlations* (or *is NC*) if $\eta_i \downarrow \eta_j$ (where η is the random string) whenever $i \neq j$.

We are primarily concerned with *conditional* negative association: $\mu \in \mathcal{M}$ is *conditionally negatively associated* (CNA) if any measure obtained from μ by conditioning on the values of some of the variables is NA. (Throughout the paper we assume that any conditioning event we consider has positive probability.) *Conditional negative correlation* (CNC) for μ is defined similarly.

When $\Omega = \{0, 1\}^n$, stronger properties are obtained by demanding NC (resp. NA) for every measure $W \circ \mu \in \mathcal{M}$ of the form

$$W \circ \mu(\eta) \propto \mu(\eta) \prod W_i^{\eta_i}$$

with $W = (W_1, \dots, W_n) \in \mathfrak{R}_+^n$. (Borrowing Ising terminology, one says that $W \circ \mu$ is obtained from μ by *imposing an external field*.) Then μ is said to be *Rayleigh* or *NC+* (resp. *NA+*), the reference in the former case being to Rayleigh's monotonicity law for electric networks (see e.g. [4] or [3]).

The competing urns model was explored in some detail by Dubhashi and Ranjan [5]*, who proved *inter alia* that threshold urn measures are NA. Another proof of this is given in [15]. Actually the argument of [5], which proves the stronger statement that the (law of the) r.v.'s

$$\xi_{ij} = \mathbf{1}_{\{\sigma(i)=j\}} \quad (3)$$

is NA, does not require identical balls. (The argument of [15] does not work for nonidentical balls.) The main purpose of the present note is to prove

Theorem 1 *Threshold urn measures are CNA.*

In contrast, as observed in [10], even ordinary urn measures need not be Rayleigh (but see the remark on R^+ in Section 5). We don't know whether Theorem 1 extends to nonidentical balls (again see Section 5).

Let us quickly say what Theorem 1 has to do with [10]. Following [15], we say that $\mu \in \mathcal{M}(\{0, 1\}^n)$ is *ultra-log-concave* (ULC) if its *rank sequence*, $\{r_i := \mu(|\eta| = i)\}_{i=0}^n$ (where $|\eta| = \sum \eta_i$), has no internal zeros and the sequence $\{r_i / \binom{n}{i}\}_{i=0}^n$ is log-concave. A set of four conjectures from [15] (see his Conjecture 4) states that each of CNC, CNA, NC+ and NA+ implies ULC; but, as shown in [1] and [10], even the weakest of these (NA+ \Rightarrow ULC) is false. Theorem 1 provides a more natural counterexample to the stronger “CNA \Rightarrow ULC,” since, as observed in [10] (disproving another conjecture from [15]), urn measures need not be ULC.

*They say “bins” rather than “urns.”

The proof of Theorem 1 gives something a little more general, as follows. Suppose that for each $j \in [n]$ we are given a sequence $0 = a_0(j) < \dots < a_{k_j}(j) = m + 1$, and for $\sigma : [m] \rightarrow [n]$ set

$$\mathbf{x}_j(\sigma) = t \text{ iff } a_t(j) \leq B_j < a_{t+1}(j). \quad (4)$$

Theorem 2 *If the $\sigma(i)$'s are i.i.d. then the \mathbf{x}_j 's in (4) are CNA.*

Call the law of $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ as in (4) a (generalized) *interval urn measure*.

The paper is organized as follows. Section 2 reduces Theorem 2 to either of our two main inequalities, (10) and (14). Each of these is valid at the level of generalized interval urns; they are equivalent in the case of ordinary urns but not obviously so in general (though the argument in Section 3 uses some interplay between the two). It is only in the derivation of Theorem 2 from (10) that we need the $\sigma(i)$'s to be i.i.d.

We give two quite different ways of getting at these main inequalities. Theorem 4 in Section 3 essentially restates (10) and (14) in induction-friendly form; the proof of the theorem given in this section is inspired by [5]. Section 4 takes a different approach, based on a graph-theoretic observation, Lemma 8, that is thought to be of independent interest. The lemma is used to: reprove (10); in combination with a result from [11] (Theorem 11 below), to prove a stronger, ultra-log-concavity version of (14); and to prove “log-submodularity” for some classes of measures.

Finally, Section 5 contains some discussion of the question of whether Theorem 1 extends to nonidentical balls, mentions a conjecture of G. Farr and a stronger one of D. Welsh, and sketches a counterexample to the latter.

Some notation. For a nonnegative vector

$$\gamma = (\gamma_{ij} : i \in [m], j \in [n]), \quad (5)$$

$A \subseteq [m]$ and $K \subseteq [n]$, the probability measure on K^A (functions from A to K) corresponding to γ is that given by

$$\Pr(\sigma) \propto W(\sigma) := \prod_{i \in A} \gamma_{i, \sigma(i)}. \quad (6)$$

Thus the r.v.'s $\sigma(i)$ are independent; they are i.i.d. if γ_{ij} does not depend on i , in which case we write simply γ_j . We also use \Pr^L ($L \subseteq [m]$) for the measure on $[n]^L$ corresponding to γ (so $\Pr = \Pr^{[m]}$).

2 Setting up

Let the law of $\sigma \in [n]^{[m]}$ be given by (6) and let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be as in (4). Let $I \cup J \cup K$ be a partition of $[n]$ and $t_j \in \{0, \dots, k_j - 1\}$ for $j \in K$, and set

$$Q := \{\mathbf{x}(\sigma) \equiv t \text{ on } K\} \quad (= \{a_{t_j}(j) \leq |\sigma^{-1}(j)| < a_{t_j+1}(j) \ \forall j \in K\}), \quad (7)$$

$X = |\sigma^{-1}(I)|$ and $Y = |\sigma^{-1}(J)|$. The main point for the proof of Theorem 2 is

$$X \downarrow Y \text{ given } Q, \quad (8)$$

given which we finish easily:

Proof of Theorem 2. With notation as above, let $\mathcal{A}, \mathcal{B} \subseteq \Omega$ be increasing events determined by I and J (more precisely, by the values of the variables \mathbf{x}_j indexed by I and J) respectively. For

Theorem 2 we should show $\mathcal{A} \downarrow \mathcal{B}$ given Q . Define $f, g : \mathbf{N} \rightarrow \mathfrak{R}$ (where $\mathbf{N} = \{0, 1, \dots\}$) by $f(k) = \Pr(\mathcal{A}|X = k)$, $g(l) = \Pr(\mathcal{B}|Y = l)$. A standard coupling argument shows that f and g are increasing, whence, according to (8),

$$\Pr(\mathcal{A} \cap \mathcal{B}|Q) = \mathbb{E}[f(X)g(Y)|Q] \leq \mathbb{E}[f(X)|Q]\mathbb{E}[g(Y)|Q] = \Pr(\mathcal{A}|Q)\Pr(\mathcal{B}|Q)$$

(where the first equality follows from conditional independence of \mathcal{A} and \mathcal{B} given (X, Y)). ■

We continue to condition on Q and write μ_k for the law of Y given $\{X = k\}$; that is,

$$\mu_k(l) = \Pr(Y = l|X = k). \quad (9)$$

We will actually prove

$$\frac{\mu_{k+1}(l+1)}{\mu_{k+1}(l)} \leq \frac{\mu_k(l+1)}{\mu_k(l)} \quad (10)$$

(whenever neither side is $0/0$, where we agree that $x/0 = \infty$ for $x > 0$), which is a strengthening of (8) once we rule out some pathologies. We recall the standard

Definition 3 $\mathcal{C} \subseteq \mathbf{N}^n$ is convex if $a, c \in \mathcal{C}$ and $a \leq b \leq c$ imply $b \in \mathcal{C}$.

It will follow from Proposition 6 below that

$$\text{supp}(\Pr) := \{(k, l) : \Pr(X = k, Y = l) > 0\} \text{ is convex.} \quad (11)$$

Given this, (10) implies that Y is stochastically decreasing in X —that is,

$$\mu_{k+1}(Y \geq t) \leq \mu_k(Y \geq t) \quad \forall k, t$$

(the easy implication is essentially Proposition 1.2 of [15])—which in turn easily implies $X \downarrow Y$.

Let

$$Z = |\sigma^{-1}(I \cup J)|. \quad (12)$$

When the $\sigma(i)$'s are i.i.d., an alternate way to specify X and Y is: let Z be as in (12), $X \sim \text{Bin}(Z, \alpha)$ and $Y = Z - X$, where $\alpha = \gamma_I / \gamma_{I \cup J}$ (with $\gamma_I = \sum_{i \in I} \gamma_i$) and $\text{Bin}(Z, \alpha)$ is the binomial distribution with parameters Z and α .

In general, for ν the law of an \mathbf{N} -valued r.v. Z and $\alpha \in [0, 1]$, let $X = X_{\nu, \alpha} \sim \text{Bin}(Z, \alpha)$, $Y = Y_{\nu, \alpha} = Z - X$ and, for lack of a better name, say ν is *binomially negatively associated* (BNA) if $X \downarrow Y$ for every α . Call a nonnegative sequence $a = (a_i)_{i=0}^\infty$ *strongly log-concave* (SLC) if

$$ia_i^2 \geq (i+1)a_{i-1}a_{i+1} \quad \forall i \geq 1 \quad (13)$$

(that is, $(i!a_i)_{i=0}^\infty$ is log-concave), and say $\nu \in \mathcal{M}(\mathbf{N})$ is SLC if the sequence $(\nu(i))_{i=0}^\infty$ is. A straightforward calculation shows that this is equivalent to saying that (10) holds for any α , $X = X_{\nu, \alpha}$ and $Y = Y_{\nu, \alpha}$ (and μ_k as in (9)): since

$$\mu_k(l) = \frac{\nu(k+l) \Pr(X = k|Z = k+l)}{\Pr(X = k)} = \frac{\nu(k+l) \binom{k+l}{k} \alpha^k (1-\alpha)^l}{\Pr(X = k)},$$

we may rewrite (10) as

$$\nu(k+l+2) \binom{k+l+2}{k+1} \nu(k+l) \binom{k+l}{k} \leq \nu(k+l+1) \binom{k+l+1}{k+1} \nu(k+l+1) \binom{k+l+1}{k},$$

which is SLC for ν . (If ν is Poisson—that is, if (13) holds with equality—then X and Y are independent Poisson r.v.’s and the inequalities (1) are equalities.) Thus, in the i.i.d. case, (10) is equivalent to saying that Z as in (12) is SLC. The latter again turns out to be true at the level of *generalized urns*; that is, for any γ as in (5), $\sigma \in [n]^{[m]}$ with law given by (6), Q as in (7) and Z as in (12),

$$\text{the law of } Z \text{ is SLC.} \quad (14)$$

It’s also easy to see that absence of internal zeros in $(\nu(i))$ is equivalent to (11) for $X = X_{\nu,\alpha}$, $Y = Y_{\nu,\alpha}$ (which, again, is given by Proposition 6), so that (14) again implies (8) (and Theorem 2). It seems interesting that both (10) and (14) are valid for generalized urns, though the equivalence that holds for i.i.d. balls disappears in the more general setting.

As mentioned earlier, in Section 4 we will combine Lemma 8 with a result from [11] to obtain an improvement of (14):

$$\text{the law of } Z \text{ is ultra-log-concave.} \quad (15)$$

(Recall—see following Theorem 1—this means that the sequence $\{\Pr(|Z| = i)/\binom{m}{i}\}_{i=0}^m$ is log-concave without internal zeros.)

3 First proof

Let \Pr be the measure on $[n]^{[m]}$ corresponding to some γ (see the end of Section 1), and for $a, b \in \mathbf{N}^{n-1}$ and $k \in \mathbf{N}$, set

$$p(k, a, b) = \Pr(B_n = k | B_j \in [a_j, b_j] \ \forall j \in [n-1]) \quad (16)$$

(recalling that $B_j = |\sigma^{-1}(j)|$).

Theorem 4 *With notation as above,*

- (a) $\frac{p(k+1, a, b)}{p(k, a, b)}$ is nonincreasing in (a, b) , and
- (b) $\frac{p(k+1, a, b)}{p(k, a, b)} \leq \frac{k}{k+1} \cdot \frac{p(k, a, b)}{p(k-1, a, b)},$

where we say nothing about the case $0/0$ and agree that $x/0 = \infty$ when $x > 0$.

See also Theorem 14 in Section 5 for a related result.

As noted earlier, part (b) of Theorem 4 is just a reformulation of (14), while (a) is a mild generalization of (10) (which in fact—see (19)—quickly reduces to (10)). To see this, note that in (10) we may assume that each of I, J is a singleton, say $K = [n-2]$, $I = \{n-1\}$, $J = \{n\}$ —formally we could pass to

$$\gamma'_{ij} = \begin{cases} \gamma_{ij} & \text{if } j \in K = [n-2] \\ \sum\{\gamma_{ij} : j \in I\} & \text{if } j = n-1 \\ \sum\{\gamma_{ij} : j \in J\} & \text{if } j = n \end{cases}$$

—and similarly in (14) we may assume $K = [n-1]$, $I = \{n\}$ and $J = \emptyset$. Then Theorem 4(b), which may also be stated

$$\text{for fixed } a, b \in \mathbf{N}^{[n-1]} \text{ the sequence } \{p(k, a, b)\} \text{ is SLC,}$$

is (up to some name changes) the same as (14), while (10) is equivalent to

$$p(k+1, a, b)/p(k, a, b) \geq p(k+1, a', b')/p(k, a', b'),$$

where $a_{n-1} = b_{n-1} = t$, $a'_{n-1} = b'_{n-1} = t+1$ (for some t) and $a_j = a'_j$, $b_j = b'_j$ for $j \in [n-2]$.

On the other hand, the inductive proof of Theorem 4 employs both the more general form of (a) and some interplay between the two parts.

Before proving the theorem we note one further consequence and give the promised Proposition 6. For $f, a \in \mathbf{N}^n$, let $\mathcal{M}_f(a) = \{\sigma \in [n]^{[m]} : |\sigma^{-1}(j)| \in [a_j, a_j + f_j] \ \forall j \in [n]\}$ and $M_f(a) = \Pr(\mathcal{M}_f(a))$. Though we won't use the next result (but see the remark following Corollary 9), it seems natural and worth mentioning.

Corollary 5 *For each $f \in \mathbf{N}^n$, $M = M_f$ satisfies the negative lattice condition:*

$$M(a)M(c) \geq M(a \vee c)M(a \wedge c) \quad \forall a, c \in \mathbf{N}^n. \quad (17)$$

This is more or less immediate from Theorem 4 once we have the next little observation, which, as noted earlier, also gives (11) and absence of internal zeros in the law of Z in (12).

Proposition 6 *For any f and M_f as above, the support of $M = M_f$ is convex.*

Proof. This will follow easily from

Claim. For any $\sigma, \tau \in [n]^{[m]}$ with $\Pr(\sigma), \Pr(\tau) > 0$ and $i \in [n]$ with $|\sigma^{-1}(i)| > |\tau^{-1}(i)|$, there are $j \in [n]$ and $\rho \in [n]^{[m]}$ with $\Pr(\rho) > 0$, $|\sigma^{-1}(j)| < |\tau^{-1}(j)|$ and

$$|\rho^{-1}(k)| = \begin{cases} |\sigma^{-1}(i)| - 1 & \text{if } k = i \\ |\sigma^{-1}(j)| + 1 & \text{if } k = j \\ |\sigma^{-1}(k)| & \text{if } k \in [n] \setminus \{i, j\}. \end{cases}$$

This is a standard type of graph-theoretic observation: regarding σ and τ as edge sets of bipartite graphs on $[m] \cup [n]$ in the natural way,[†] we need a path with edges alternately from $\sigma \setminus \tau$ and $\tau \setminus \sigma$ that begins with a σ -edge at i and ends with a τ -edge at some j as above. (We then get ρ by switching σ and τ on this path.) We omit the routine proof that such a path must exist.

To prove Proposition 6, we should show that for all distinct $a, b, c \in \mathbf{N}^n$ with $a \leq b \leq c$ and $a, c \in \text{supp}(M)$, we also have $b \in \text{supp}(M)$. Of course it suffices to show this when there is some $i \in [n]$ with $b_i = c_i - 1$ and $b_k = c_k$ for all $k \neq i$. Choose $\tau \in \mathcal{M}(a) := \mathcal{M}_f(a)$ and $\sigma \in \mathcal{M}(c)$ with $\Pr(\tau), \Pr(\sigma) > 0$. We assume $|\sigma^{-1}(i)| = c_i + f_i$, since otherwise $\sigma \in \mathcal{M}(b)$ and we are finished. Letting j, ρ be as in the claim (note $|\tau^{-1}(i)| < c_i + f_i$), we have

$$|\rho^{-1}(i)| = b_i + f_i$$

and

$$|\rho^{-1}(j)| = |\sigma^{-1}(j)| + 1 \in [c_j + 1, a_j + f_j] \subseteq [b_j, b_j + f_j],$$

whence $\rho \in \mathcal{M}(b)$ and $b \in \text{supp}(M)$. ■

[†]We pretend $[m] \cap [n] = \emptyset$.

Proof of Corollary 5. It is easy to see (and standard) that convexity of M (given by Proposition 6) implies that it's enough to prove (17) when there are indices i and j with $a_i = c_i - 1$, $a_j = c_j + 1$, and $a_k = c_k$ for all $k \neq i, j$. In this case—assuming, w.l.o.g., that $i = n - 1$ and $j = n$ —we set

$$p_1(k) = \Pr(B_n = k | B_l \in [a_l, a_l + f_l] \ \forall l \in [n - 1])$$

and

$$p_2(k) = \Pr(B_n = k | B_l \in [c_l, c_l + f_l] \ \forall l \in [n - 1]).$$

Then (17) is

$$\left(\sum_{k=c_n+1}^{c_n+f_n+1} p_1(k) \right) \left(\sum_{k=c_n}^{c_n+f_n} p_2(k) \right) \geq \left(\sum_{k=c_n+1}^{c_n+f_n+1} p_2(k) \right) \left(\sum_{k=c_n}^{c_n+f_n} p_1(k) \right)$$

and follows immediately from

$$p_1(k)p_2(l) \geq p_1(l)p_2(k) \text{ whenever } k \geq l,$$

which is a consequence of Theorem 4(a) (and Proposition 6). ■

We now assume (as we may) that $\sum_j \gamma_{ij} = 1$ for each i . The proof of Theorem 4 resembles that of Theorem 33 in [5], and is based on

Observation 7 *For any $i \in [n]$, $k \in \mathbf{N}$, and event Q determined by $(\sigma^{-1}(j) : j \neq i)$,*

$$\Pr(B_i = k + 1, Q) = \frac{1}{k + 1} \sum_{l \in [m]} \gamma_{li} \Pr^{[m] \setminus \{l\}}(B_i = k, Q).$$

(Recall \Pr^L was defined at the end of Section 1.) We also use the trivial

$$\min_i \frac{\alpha_i}{\beta_i} \leq \frac{\alpha_1 + \cdots + \alpha_k}{\beta_1 + \cdots + \beta_k} \leq \max_i \frac{\alpha_i}{\beta_i} \quad (18)$$

(for all $\alpha_1, \dots, \alpha_k, \beta_1, \dots, \beta_k \geq 0$ with $\beta_1 + \cdots + \beta_k > 0$, where, again, $x/0 := \infty$ when $x > 0$).

Proof of Theorem 4. We proceed by induction on m , omitting the easy base cases with $m = 1$. For (a), it's enough to show that the ratio in question does not increase when we increase a single entry—w.l.o.g. the $(n - 1)$ st—of one of a, b . Thus, by (18), it suffices to show that

$$\frac{\Pr(B_n = k + 1 | B_{n-1} = t, R)}{\Pr(B_n = k | B_{n-1} = t, R)} \text{ is nonincreasing in } t,$$

where $R = \{a_j \leq B_j \leq b_j \ \forall j \in [n - 2]\}$; and by Proposition 6, this will follow if we show

$$\frac{\Pr(B_n = k + 1 | B_{n-1} = t + 1, R)}{\Pr(B_n = k | B_{n-1} = t + 1, R)} \leq \frac{\Pr(B_n = k + 1 | B_{n-1} = t, R)}{\Pr(B_n = k | B_{n-1} = t, R)} \quad (19)$$

for all t for which the probabilities appearing in (19) are positive. (This is the easy reduction of (a) to (10) mentioned earlier.)

By Observation 7 we may write the left side of (19) as

$$\frac{\sum_{l \in [m]} \gamma_{l,n-1} \Pr^{[m] \setminus \{l\}}(B_n = k+1, B_{n-1} = t, R)}{\sum_{l \in [m]} \gamma_{l,n-1} \Pr^{[m] \setminus \{l\}}(B_n = k, B_{n-1} = t, R)},$$

which, by (18), is at most

$$\max_{l \in [m]} \frac{\Pr^{[m] \setminus \{l\}}(B_n = k+1 | B_{n-1} = t, R)}{\Pr^{[m] \setminus \{l\}}(B_n = k | B_{n-1} = t, R)}.$$

Thus, setting $Q = \{B_{n-1} = t\} \wedge R$ and assuming (w.l.o.g.) that the maximum occurs at $l = m$, we will have (19) if we show

$$\frac{\Pr(B_n = k+1 | Q)}{\Pr(B_n = k | Q)} \geq \frac{\Pr^{[m-1]}(B_n = k+1 | Q)}{\Pr^{[m-1]}(B_n = k | Q)}. \quad (20)$$

Now

$$\frac{\Pr(B_n = k+1 | Q)}{\Pr(B_n = k | Q)} = \frac{\sum_{j \in [n]} \Pr(\sigma(m) = j | Q) \Pr(B_n = k+1 | Q, \sigma(m) = j)}{\sum_{j \in [n]} \Pr(\sigma(m) = j | Q) \Pr(B_n = k | Q, \sigma(m) = j)},$$

so that (20) will follow (again using (18)) from

$$\frac{\Pr(B_n = k+1 | Q, \sigma(m) = j)}{\Pr(B_n = k | Q, \sigma(m) = j)} \geq \frac{\Pr^{[m-1]}(B_n = k+1 | Q)}{\Pr^{[m-1]}(B_n = k | Q)} \quad \text{for all } j \in [n] \quad (21)$$

(where, again, “for all $j \in [n]$ ” really includes only those for which $\Pr(Q, \sigma(m) = j) > 0$).

There are three cases to consider. If $j = n$, the left side of (21) is

$$\frac{\Pr^{[m-1]}(B_n = k | Q)}{\Pr^{[m-1]}(B_n = k-1 | Q)},$$

which is at least the right side of (21) by (part (b) of) our induction hypothesis. If $j = n-1$, the left side of (21) is

$$\frac{\Pr^{[m-1]}(B_n = k+1 | B_{n-1} = t-1, R)}{\Pr^{[m-1]}(B_n = k | B_{n-1} = t-1, R)},$$

which is at least the right side of (21) by (part (a) of) the induction hypothesis. Finally, if $j \neq n-1, n$, the left side of (21) is

$$\frac{\Pr^{[m-1]}(B_n = k+1 | B_{n-1} = t, R^*)}{\Pr^{[m-1]}(B_n = k | B_{n-1} = t, R^*)},$$

where R^* is obtained from R by replacing the condition $a_j \leq B_j \leq b_j$ by the condition $a_j - 1 \leq B_j \leq b_j - 1$; again this is at least the right side of (21) by part (a) of the induction hypothesis.

We now turn to (b) and set $Q = \{a_j \leq B_j \leq b_j \mid \forall j \in [n-1]\}$. Then we have, again using Observation 7 and (18),

$$\begin{aligned} \frac{\Pr(B_n = k+1 | Q)}{\Pr(B_n = k | Q)} &= \frac{k \sum_{l \in [m]} \gamma_{ln} \Pr^{[m] \setminus \{l\}}(B_n = k, Q)}{(k+1) \sum_{l \in [m]} \gamma_{ln} \Pr^{[m] \setminus \{l\}}(B_n = k-1, Q)} \\ &\leq \max_{l \in [m]} \frac{k \Pr^{[m] \setminus \{l\}}(B_n = k, Q)}{(k+1) \Pr^{[m] \setminus \{l\}}(B_n = k-1, Q)} \\ &\stackrel{\text{w.l.o.g.}}{=} \frac{k \Pr^{[m-1]}(B_n = k, Q)}{(k+1) \Pr^{[m-1]}(B_n = k-1, Q)} \end{aligned}$$

(noting that we may assume, by Proposition 6, that $\Pr(B_n = r|Q)$ is positive for $r \in \{k-1, k, k+1\}$); so we will be done if we can show

$$\frac{\Pr^{[m-1]}(B_n = k|Q)}{\Pr^{[m-1]}(B_n = k-1|Q)} \leq \frac{\Pr(B_n = k|Q)}{\Pr(B_n = k-1|Q)}. \quad (22)$$

Proceeding as in the proof of part (a), we may rewrite

$$\frac{\Pr(B_n = k|Q)}{\Pr(B_n = k-1|Q)} = \frac{\sum_{j \in [n]} \Pr(\sigma(m) = j|Q) \Pr(B_n = k|Q, \sigma(m) = j)}{\sum_{j \in [n]} \Pr(\sigma(m) = j|Q) \Pr(B_n = k-1|Q, \sigma(m) = j)},$$

so for (22) it is enough to show that, for each $j \in [n]$,

$$\frac{\Pr(B_n = k|Q, \sigma(m) = j)}{\Pr(B_n = k-1|Q, \sigma(m) = j)} \geq \frac{\Pr^{[m-1]}(B_n = k|Q)}{\Pr^{[m-1]}(B_n = k-1|Q)},$$

which, as did (21), follows easily from our induction hypothesis (here we only need to consider the two cases $j = n$ and $j \neq n$). ■

4 A graphical approach

We begin here with a natural and seemingly new graph theoretic statement which we regard as the main point of this section. Given a multigraph G on vertex set V and $a, b \in \mathbf{N}^V$, let $\mathcal{O}(a, b) = \mathcal{O}_G(a, b)$ be the set of orientations of G for which

$$d^+(x) \geq a_x \quad \text{and} \quad d^-(x) \geq b_x \quad \text{for all } x \in V$$

and $N(a, b) = N_G(a, b) = |\mathcal{O}(a, b)|$. Here d^+ and d^- are, as usual, out- and in-degrees. We will also use d_x for the degree of x in G . Note we regard a loop (at x , say) as having two orientations, each of which contributes 1 to each of $d^+(x)$ and $d^-(x)$.

Lemma 8 *If $a, b, r, s \in \mathbf{N}^V$ satisfy*

$$a \geq r, s \quad \text{and} \quad a + b \geq r + s$$

(where the inequalities are with respect to the product order on \mathbf{N}^V), then

$$N(a, b) \leq N(r, s). \quad (23)$$

Of course the idea is that it's harder to satisfy a set of demands that always requires large out-degrees than one for which these requirements are mixed. For the sake of comparison, let us also mention the specialization of Corollary 5 to the present situation:

Corollary 9 *If $a + b = r + s$ then*

$$N(a, b)N(r, s) \geq N(a \vee r, b \wedge s)N(a \wedge r, b \vee s). \quad (24)$$

Proof. Interpret vertices of G as urns and edges as balls, and assume that for each edge (ball) e we have $\gamma_{ex} = 1$ or 0 according to whether x is or is not an end of e . Then (24) is just Corollary 5 with $f = d - a - b (= d - r - s)$, where $d = (d_x : x \in V(G))$ is the vector of degrees.

■

Remark. It's possible to simplify the proof of Lemma 8 using Corollary 9; but of course this depends on Theorem 4, so is really harder than the following direct proof. On the other hand, it's not too hard to derive Theorem 4(a) from Lemma 8; see [14]. (And below we use Lemma 8 to prove (15), which is stronger than Theorem 4(b).)

Proof of Lemma 8. We proceed by induction on $\varphi(G, a, b) := |E(G)| + \sum_{x \in V} (d_x - a_x - b_x)$, calling $x \in V$ *saturated* if $a_x + b_x = d_x$. Since N is nonincreasing in each of its arguments, we may assume $a + b = r + s$ (or we can increase r or s).

Suppose first that there is at least one saturated vertex, x . We may assume there are no loops at x , since otherwise (23) follows easily from the induction hypothesis applied to the graph gotten from G by deleting such loops. Let $\alpha = a_x, \beta = b_x, \rho = r_x, \sigma = s_x$, and let $X = \{e_1, \dots, e_{\alpha+\beta}\}$ be the set of edges incident with x .

Consider a set π consisting of β pairs $\{e_i, e_j\} \subseteq X$, with the 2β edges appearing in π distinct, and, say, y_i the vertex joined to x by e_i (so the y_i 's need not be distinct). Let $G(\pi)$ be the graph with vertex set $V \setminus \{x\}$ and edge set $E(G) \setminus X \cup \{e_{ij} : \{e_i, e_j\} \in \pi\}$, where e_{ij} joins y_i and y_j . Let $U(\pi)$ be the set of edges in X not belonging to pairs from π , and $U_z(\pi)$ the set of edges of $U(\pi)$ incident to z .

Define $a^\pi, b^\pi \in \mathbf{N}^{V(G(\pi))}$ by

$$a_z^\pi = a_z \text{ and } b_z^\pi = \max\{b_z - |U_z(\pi)|, 0\} \text{ for all } z \in V \setminus \{x\} \quad (= V(G(\pi))).$$

For each π as above and $T \in \binom{U(\pi)}{\rho-\beta}$ (where $\binom{A}{k} = \{B \subseteq A : |B| = k\}$), define $r^{\pi,T}, s^{\pi,T} \in \mathbf{N}^{V(G(\pi))}$ by

$$r_z^{\pi,T} = \max\{r_z - |U_z(\pi) \setminus T|, 0\}$$

and

$$s_z^{\pi,T} = \max\{s_z - |U_z(\pi) \cap T|, 0\}$$

for all $z \in V \setminus \{x\}$.

Each $\sigma \in \mathcal{O}_{G(\pi)}(a^\pi, b^\pi)$ maps naturally to a (unique) $\hat{\sigma} \in \mathcal{O}_G(a, b)$, namely: $\hat{\sigma}$ agrees with σ on $E(G - x)$; orients all edges of $U(\pi)$ away from x ; and orients e_i from y_i to x and e_j from x to y_j whenever σ orients e_{ij} from y_i to y_j (where, when $y_i = y_j$, we interpret one orientation of the loop e_{ij} as $y_i \rightarrow y_j$ and the other as $y_j \rightarrow y_i$). Since each $\tau \in \mathcal{O}_G(a, b)$ is in the range of this map for exactly $\binom{\alpha}{\beta} \beta!$ choice of π , we have

$$N(a, b) = \frac{1}{\binom{\alpha}{\beta} \beta!} \sum_{\pi} N_{G(\pi)}(a^\pi, b^\pi). \quad (25)$$

Similarly,

$$N(r, s) = \frac{1}{\binom{\rho}{\beta} \binom{\sigma}{\beta} \beta!} \sum_{\pi} \sum_{T \in \binom{U(\pi)}{\rho-\beta}} N_{G(\pi)}(r^{\pi,T}, s^{\pi,T}). \quad (26)$$

Since $a^\pi + b^\pi \geq r^{\pi,T} + s^{\pi,T}$ and $a^\pi \geq r^{\pi,T}, s^{\pi,T}$, it follows from the induction hypothesis that

$$N_{G(\pi)}(a^\pi, b^\pi) \leq N_{G(\pi)}(r^{\pi,T}, s^{\pi,T}) \text{ for all } \pi \text{ and } T \in \binom{U(\pi)}{\rho-\beta}. \quad (27)$$

(Note that $\varphi(G(\pi), a^\pi, b^\pi) < \varphi(G, a, b)$, since $|E(G(\pi))| < |E(G)|$ and, for $z \in V \setminus \{x\}$, $a_z^\pi + b_z^\pi \geq a_z + b_z - |U_z(\pi)|$, while the degree of z in $G(\pi)$ is $d_z - |U_z(\pi)|$.) Combining (25), (26) and (27), we have

$$N(r, s) \geq \frac{1}{\binom{\rho}{\beta} \binom{\sigma}{\beta} \beta!} \sum_{\pi} \binom{\alpha - \beta}{\rho - \beta} N_{G(\pi)}(a^\pi, b^\pi) = \frac{\alpha! \beta!}{\rho! \sigma!} N(a, b) \geq N(a, b),$$

where the last inequality follows from the assumptions $\alpha \geq \rho, \sigma$ and $\alpha + \beta = \rho + \sigma$.

So we may assume there are no saturated vertices. In this case we fix $x \in V$ with $a_x > b_x$. (Of course if there is no such vertex, then $a = b = r = s$ and (23) is an equality.) For $\gamma, \delta \in \mathbf{N}$ let $N'(\gamma, \delta)$ be the number of orientations of G with

$$(d_y^+, d_y^-) \geq \begin{cases} (a_y, b_y) & \text{if } y \neq x \\ (\gamma, \delta) & \text{if } y = x, \end{cases}$$

and let $N''(\gamma, \delta)$ be defined analogously with (r, s) in place of (a, b) . Let $\alpha = a_x, \beta = b_x, \rho = r_x, \sigma = s_x$, so that (23) is

$$N'(\alpha, \beta) \leq N''(\rho, \sigma). \quad (28)$$

By induction we have

$$N'(\gamma, \delta) \leq N''(\eta, \xi) \text{ whenever } \gamma \geq \eta, \delta \geq \xi \text{ and } \gamma + \delta = \eta + \xi > \alpha + \beta. \quad (29)$$

We apply this to the identity

$$N'(\alpha, \beta) = N'(\alpha, \beta + 1) + N'(d_x - \beta, \beta). \quad (30)$$

If $\alpha > \sigma$, then, by (29), the right side of (30) is at most

$$N''(\rho, \sigma + 1) + N''(d_x - \sigma, \sigma) = N''(\rho, \sigma).$$

If $\alpha > \rho$, then, again using (29), the right side of (30) is at most

$$N''(\rho + 1, \sigma) + N''(\rho, d_x - \rho) = N''(\rho, \sigma).$$

(And, since $\alpha > \beta$, we have at least one of $\alpha > \sigma, \alpha > \rho$.)

■

The next result isolates (and generalizes) the main point in the derivation of (15) from Lemma 8. We consider a hypergraph $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$ on a set W of size $2l$, where

- (i) the edges of \mathcal{H}_1 are pairwise disjoint and
- (ii) the edges of \mathcal{H}_2 are of size 2 and pairwise disjoint.

Let S be the set of vertices of \mathcal{H} not covered by edges of \mathcal{H}_2 , and $|S| = 2t$. Given $\alpha : \mathcal{H}_1 \rightarrow \mathbf{N}$, let N_i be the number of partitions (X, Y) of W with each of X, Y a vertex cover of \mathcal{H}_2 , each of $|X \cap H|, |Y \cap H|$ at least α_H for each $H \in \mathcal{H}_1$, and $|X| = i$.

Lemma 10 *In the above situation, $N_l \geq \frac{t+1}{t} N_{l+1}$.*

Proof. For $i \in \mathbf{N}$ and π a collection of $t - 1$ disjoint 2-sets contained in S , let $\mathcal{N}_i(\pi)$ be the set of partitions as above for which each of X, Y also covers the edges of π , and set $N_i = |\mathcal{N}_i|$. We assert that (for each π)

$$N_l(\pi) \geq 2N_{l+1}(\pi). \quad (31)$$

This implies the proposition since (as is easily seen)

$$N_l = \frac{1}{t \cdot t!} \sum_{\pi} N_l(\pi)$$

and

$$N_{l+1} = \frac{1}{\binom{t+1}{2}(t-1)!} \sum_{\pi} N_{l+1}(\pi).$$

For the proof of (31) let x, y be the two vertices of W not contained in members of $\mathcal{H}'_2 := \mathcal{H}_2 \cup \pi$. Noting that $(X, Y) \in \mathcal{N}_{l+1}(\pi)$ implies $\{x, y\} \in X$, we may regard (X, Y) as an orientation of \mathcal{H}'_2 , where orienting $\{u, v\}$ from u to v corresponds to putting u in X (and v in Y). The orientations corresponding to (X, Y) 's from \mathcal{N}_{l+1} are those for which, for each $H \in \mathcal{H}_1$,

$$d^+(H) \geq \alpha_H - |H \cap \{x, y\}| \quad \text{and} \quad d^-(H) \geq \alpha_H,$$

where, for the given orientation, $d^+(H)$ (resp. $d^-(H)$) is the number of oriented edges whose tails (resp. heads) lie in H .

If we let G be the multigraph gotten from (\mathcal{H}, π) by collapsing each $H \in \mathcal{H}_1$ to a single vertex (so for example, any $\{u, v\} \in \mathcal{H}'_2$ contained in some $H \in \mathcal{H}_1$ becomes a loop in G), then the above discussion says that $N_{l+1}(\pi) = N_G(a, b)$ (see Lemma 8 for the notation), where

$$a_z = \alpha_H - |H \cap \{x, y\}| \quad \text{and} \quad b_z = \alpha_H$$

if z is the vertex of G corresponding to $H \in \mathcal{H}_1$, and $a_z = b_z = 0$ if z is not of this type (i.e. $z \in W \setminus \cup\{H : H \in \mathcal{H}_1\}$).

A similar discussion shows that $N_l(\pi) = N_G(r, s) + N_G(s, r)$, where

$$r_z = \alpha_H - \mathbf{1}_{\{x \in H\}} \quad \text{and} \quad s_z = \alpha_H - \mathbf{1}_{\{y \in H\}}$$

if z is the vertex of G corresponding to $H \in \mathcal{H}_1$, and $r_z = s_z = 0$ if z is not of this type. (For example, $N_G(r, s)$ counts pairs (X, Y) with $x \in X$ (and $y \in Y$).)

Finally, Lemma 8 gives $N_G(r, s), N_G(s, r) \geq N_G(a, b)$, so we have (31). (Strictly speaking we may be applying Lemma 8 with some negative entries in b, r and/or s ; but it's easy to see that this slightly more general version follows from the lemma as stated.)

■

As mentioned earlier, the proof of (15) also requires Theorem 11 below. (If we just wanted (14) then Lemma 10 alone would suffice.) For $\mu \in \mathcal{M}_m := \mathcal{M}(2^{[m]})$, set

$$\alpha_i(\mu) = \binom{m}{i}^{-1} \sum \{\mu(A)\mu(\bar{A}) : A \subseteq [m], |A| = i\} \quad (32)$$

(where $\bar{A} = [m] \setminus A$). Say $\mu \in \mathcal{M}_{2k}$ has the *antipodal pairs property* (APP) if $\alpha_k(\mu) \geq \alpha_{k-1}(\mu)$, and $\mu \in \mathcal{M}_m$ has the *conditional antipodal pairs property* (CAPP) if every measure obtained from μ by conditioning on the values of some $m - 2k$ variables (for some k) has the APP (where we view conditioning on the values indexed by T as producing a measure in $\mathcal{M}(2^{[m] \setminus T})$).

Theorem 11 ([11]) *A measure with the CAPP and no internal zeros in its rank sequence is ULC.*

Proof of (15). It's again enough to show this when $K = [n-1]$, $I = \{n\}$ and $J = \emptyset$. Setting $U = \sigma^{-1}(K)$ and letting μ be the law of U , we prove the equivalent

$$\mu \text{ is ULC.} \quad (33)$$

For $A \subseteq [m]$ and $\sigma \in K^A$, say $\sigma \in Q$ if it satisfies the conditions in (7), which we now rewrite

$$S_j \leq |\sigma^{-1}(j)| \leq T_j \quad \forall j \in K, \quad (34)$$

where $S_j = a_{t_j}(j)$ and $T_j = a_{t_j+1}(j) - 1$. (This extends the Q of (7), which was a subset of $[n]^{[m]}$.) Write $\sigma \sim A$ if $\sigma \in K^A \cap Q$ and $\sigma \sim l$ if $\sigma \sim A$ for some A of size l , and set $T(A) = \prod \{\gamma_{in} : i \in [m] \setminus A\}$. Then μ is given by

$$\mu(A) \propto T(A) \sum \{W(\sigma) : \sigma \sim A\} \quad (A \subseteq [m]).$$

By Theorem 11 we will have (33) if we show that μ satisfies the CAPP and its rank sequence has no internal zeros. The latter condition is given by Proposition 6, applied with

$$f_j = \begin{cases} T_j - S_j & \text{if } j \in K \\ 0 & \text{if } j = n \end{cases}$$

(so that $\mu(|U| = k) = M_f(S_1, \dots, S_{n-1}, m - k)$).

To show that μ has the CAPP, we should verify the APP for the conditional measures $\mu_{X,Y} \in \mathcal{M}_{Y \setminus X}$ given by

$$\mu_{X,Y}(A) \propto \mu(A \cup X) \quad (A \subseteq Y \setminus X),$$

where $X \subseteq Y \subseteq [m]$ and $|Y \setminus X| = 2k$ (for some k). Fixing X, Y and letting A, B run over subsets of $Z := Y \setminus X$, this amounts to

$$\begin{aligned} T(X)T(Y) \sum_{|A|=k} \sum \{W(\sigma)W(\tau) : \sigma \sim A \cup X, \tau \sim (Z \setminus A) \cup X\} \\ \geq \frac{k+1}{k} T(X)T(Y) \sum_{|B|=k-1} \sum \{W(\alpha)W(\beta) : \alpha \sim B \cup X, \beta \sim (Z \setminus B) \cup X\} \end{aligned} \quad (35)$$

Regard each of $\sigma, \tau, \alpha, \beta$ in (35) as a bipartite graph on the vertex set $[m] \cup K$ in the natural way. Then for each pair (σ, τ) appearing in (35) the *multiset* union $G = \sigma \cup \tau$ is a bipartite multigraph with exactly $2(|X| + k)$ edges and

$$d_G(i) = \begin{cases} 2 & \text{if } i \in X \\ 1 & \text{if } i \in Y \setminus X \\ 0 & \text{if } i \in [m] \setminus Y \end{cases}$$

(and similarly for pairs (α, β)). We may thus rewrite (35) as

$$\begin{aligned} \sum_G \sum \{W(\sigma)W(\tau) : \sigma \cup \tau = G, \sigma \sim |X| + k\} \\ \geq \frac{k+1}{k} \sum_G \sum \{W(\alpha)W(\beta) : \alpha \cup \beta = G, \alpha \sim |X| + k - 1\}, \end{aligned}$$

and it is enough to show that for each fixed G we have the corresponding inequality for the inner sum, i.e.

$$\sum \{W(\sigma)W(\tau) : \sigma \cup \tau = G, \sigma \sim |X|+k\} \geq \frac{k+1}{k} \sum \{W(\alpha)W(\beta) : \alpha \cup \beta = G, \alpha \sim |X|+k-1\}. \quad (36)$$

This has the advantage that the weights no longer play a role, since for σ, \dots, β as in (36),

$$W(\sigma)W(\tau) = W(\alpha)W(\beta);$$

so we will have (36) if we show

$$N_k \geq \frac{k+1}{k} N_{k-1}, \quad (37)$$

where $N_i = N_i(G)$ is the number of partitions $G = \gamma \cup \delta$ with $\gamma \sim |X| + i$.

Now let $\mathcal{H} = \mathcal{H}_1 \cup \mathcal{H}_2$ be the hypergraph with vertex set $W = E(G)$,

$$\mathcal{H}_1 = \{H_j : j \in K\},$$

where $H_j = \{e \in W : j \in e\}$, and

$$\mathcal{H}_2 = \{\{e, f\} : e \neq f, e \text{ and } f \text{ have the same end in } [m]\}.$$

Define $\alpha : \mathcal{H}_1 \rightarrow \mathbf{N}$ by

$$\alpha_{H_j} = \max\{S_j, d_G(j) - T_j\}$$

(recall S_j, T_j were defined following (34)). We are then in the situation of Lemma 10: a partition $W = W_1 \cup W_2$ with $|W_1| = |X| + i$ (and $|W_2| = |Y| - i$) as in the lemma (i.e. with (W_1, W_2) in place of (X, Y)) is the same thing as a partition $G = \gamma \cup \delta$ with $\gamma \sim |X| + i$ (and $\delta \sim |Y| - i$), and the t in the lemma is equal to the present k ; so we have (37). ■

Remark. Log-concavity results being of some interest, we mention one appealing specialization of (15). (See e.g. [16] or [2] for much more on log-concavity in combinatorial settings.) For a bipartite graph $G = (V \cup K, E)$, define a G -map to be a function $f : A \rightarrow K$ with $A \subseteq V$ and $(v, f(v)) \in E$ for all $v \in A$. Given $l, u \in \mathbf{N}^K$ (with $l_j \leq u_j$), call a G -map *valid* if $|f^{-1}(j)| \in [l_j, u_j]$ for all $j \in K$. Let $s_k = s_k(G, l, u)$ be the number of valid G -maps $f : A \rightarrow K$ with $|A| = k$.

Theorem 12 *For any G, l, u the sequence $(s_0, \dots, s_{|V|})$ is ultra-log-concave.*

In the special case $l \equiv 0, u \equiv 1$, s_k becomes $\Phi_k = \Phi_k(G)$, the number of matchings of size k in G . Heilman and Lieb [8, 9] and Kunz [12] (see also [13, Chapter 8]) proved that for any (not necessarily bipartite) graph G , the *matching generating polynomial*

$$p(x) = \sum_{k=0}^{\nu} \Phi_k x^k$$

(where, as usual, $\nu = \max\{k : \Phi_k > 0\}$ is the matching number of G) has all real (negative) roots. This implies, by Newton's inequalities (e.g. [7, Theorem 51]), that

$$(\Phi_0, \dots, \Phi_{\nu}) \text{ is ULC}, \quad (38)$$

which, if $\nu < |V|$, is somewhat stronger than this case of Theorem 12. In contrast, for general l and u the polynomial

$$\sum_{k=0}^{|V|} s_k x^k$$

need not have all real roots. (For example, let $V = \{y, v, w\}$, $K = \{1\}$, $E(G) = \{\{y, 1\}, \{v, 1\}, \{w, 1\}\}$, $l_1 = 1, u_1 = 3$.)

Actually Theorem 11 can be used to show that for any G , $(\Phi_0, \dots, \Phi_\tau)$ is ULC (where, as usual, $\tau = \tau(G)$ is the vertex cover number), which in particular recovers (38) when G is bipartite. As this doesn't use Lemma 8, we won't go into it here. It would be very interesting to see a combinatorial proof of (38) for general G .

Before closing this section we point out one further consequence of Lemma 8, which seems to us interesting for its own sake. With notation as in the above proof of (15), set $f(A) = \sum \{W(\sigma) : \sigma \in K^A \cap Q\}$ ($A \subseteq [m]$). We assert that f satisfies the negative lattice condition:

$$f(A \cup B)f(A \cap B) \leq f(A)f(B) \quad \forall A, B \subseteq [m]. \quad (39)$$

While we don't see how to get (14) (or (15)) from this in general, it's not hard to see that it does imply (14) in case the $\sigma(i)$'s are i.i.d., so gives yet another proof Theorem 2. We omit the details.

Proof of (39). We may rewrite the inequality as

$$\sum \sum \{W(\sigma)W(\tau) : \sigma \sim A \cup B, \tau \sim A \cap B\} \leq \sum \sum \{W(\alpha)W(\beta) : \alpha \sim A, \beta \sim B\}. \quad (40)$$

As before we regard $\sigma, \tau, \alpha, \beta$ in (40) as bipartite graphs on $[m] \cup K$. For each pair (σ, τ) appearing in (40), the (multiset) union $G = \sigma \cup \tau$ is a bipartite multigraph with

$$d_G(i) = \begin{cases} 2 & \text{if } i \in A \cap B \\ 1 & \text{if } i \in A \Delta B \\ 0 & \text{otherwise} \end{cases}$$

(and similarly for pairs (α, β)), and it's enough to show that, for each such G , (40) still holds if we restrict to pairs (σ, τ) and (α, β) with

$$\sigma \cup \tau = \alpha \cup \beta = G. \quad (41)$$

Again the weights ($W(\sigma)$ etc.) cancel and it's enough to show

$$N(A \cup B, A \cap B) \leq N(A, B), \quad (42)$$

where, for $C, D \subseteq [m]$, $N(C, D) = N_G(C, D)$ is the number of partitions $E(G) = \gamma \cup \delta$ with $\gamma \sim C$ and $\delta \sim D$.

Notice now that we are really counting partitions $\hat{\sigma} \cup \hat{\tau}$ and $\hat{\alpha} \cup \hat{\beta}$ of the edges of $G' := G[(A \cap B) \cup K]$, since for any σ, \dots, β (as in (40)) satisfying (41), any edge of G with an end in $A \setminus B$ (resp. $B \setminus A$) must belong to $\sigma \cap \alpha$ (resp. $\sigma \cap \beta$).

For $j \in K$ and $C \subseteq [m]$ write $d_C(j)$ for the number of edges of G joining j to C . In terms of $\hat{\sigma}, \dots, \hat{\beta}$ the requirement that σ, \dots, β satisfy (34) becomes the condition that for each $j \in K$,

$$\begin{aligned} S_j - d_{A \Delta B}(j) &\leq |\hat{\sigma}^{-1}(j)| \leq T_j - d_{A \Delta B}(j); & S_j &\leq |\hat{\tau}^{-1}(j)| \leq T_j; \\ S_j - d_{A \setminus B}(j) &\leq |\hat{\alpha}^{-1}(j)| \leq T_j - d_{A \setminus B}(j) & \text{and} & S_j - d_{B \setminus A}(j) \leq |\hat{\beta}^{-1}(j)| \leq T_j - d_{B \setminus A}(j). \end{aligned} \quad (43)$$

Now let H be the multigraph on vertex set $E(G')$ with edge set $\{e_x : x \in A \cap B\}$, where e_x joins the two edges of G' containing x . We may identify a partition $E(G') = \gamma \cup \delta$ with the orientation of H gotten by directing e_x from a to b whenever $a \in \gamma$ and $b \in \delta$, where a, b are the edges on x in G' . The orientations corresponding to pairs $(\hat{\sigma}, \hat{\tau})$ as in (43) are then those satisfying

$$S_j - d_{A \Delta B}(j) \leq d^+(j) \leq T_j - d_{A \Delta B}(j) \text{ and } S_j \leq d^-(j) \leq T_j \quad \forall j \in K$$

while those corresponding to pairs $(\hat{\alpha}, \hat{\beta})$ are those with

$$S_j - d_{A \setminus B}(j) \leq d^+(j) \leq T_j - d_{A \setminus B}(j) \text{ and } S_j - d_{B \setminus A}(j) \leq d^-(j) \leq T_j - d_{B \setminus A}(j) \quad \forall j \in K.$$

That the number of orientations of the first type is at most the number of the second type is then an instance of Lemma 8. ■

5 Final remarks

The most interesting question left open by the present work is whether Theorem 1 (even without thresholds) extends to nonidentical balls; that is,

Question 13 *Are generalized urn measures (or generalized threshold or interval urn measures) CNA?*

As mentioned in the introduction, Dubhashi and Ranjan [5] showed NA for the ξ_{ij} 's defined in (3), which immediately gives NA for generalized threshold urn measures. That the weaker (than CNA) CNC, at least, does hold for generalized threshold (or, more generally, “interval”) urn measures is a special case of the following result; this is a somewhat more general version of Corollary 34 of [5], which implies CNC for generalized threshold urn measures. We again take \Pr to be the measure on $[n]^{[m]}$ corresponding to some γ , and, for $\mathcal{A} \subseteq 2^{[m]}$ and $a, b \in \mathbf{N}^{n-1}$, set $p(\mathcal{A}, a, b) = \Pr(\sigma^{-1}(n) \in \mathcal{A} | B_j \in [a_j, b_j] \quad \forall j \in [n-1])$.

Theorem 14 *For any increasing \mathcal{A} , $p(\mathcal{A}, a, b)$ is decreasing in (a, b) .*

The proof is more or less the same as that of Theorem 4(a), so will not be given here; see [14, Theorem 1.23]. (The proof of Corollary 34 in [5] is not quite correct, since it depends on the incorrect Proposition 24.)

Thus one reason to be interested in whether generalized urn measures are CNA is that a negative answer would provide a counterexample to an important conjecture of Pemantle [15] stating that CNC implies CNA. (He also conjectures that the Rayleigh property NC+ implies NA+.) Pursuing this a little further, say $\mu \in \mathcal{M}(\{0, 1\})^n$ is R^+ if every $W \circ \mu$ with $W_i \in \{0\} \cup [1, \infty) \quad \forall i$ is NC. An easy simulation shows that CNC for the class of generalized urn measures is the same as R^+ for this class. (Note this is without thresholds; it's easy to see that R^+ need not hold for (even ordinary) threshold urn measures.) So failure of CNA here would in fact disprove

Conjecture 15 *R^+ implies CNA,*

a weakening of the first conjecture of Pemantle above.

At this writing we can (e.g.) give a positive answer to Question 13 when each ball chooses from just two urns; this is of course quite special, but seems of some interest since it corresponds

to in- and out-degree statistics for a random orientation of a graph (where edges are oriented independently, but the two orientations of an edge may have different probabilities). Even this special case seems to require an interesting argument, but we will not give this here as the paper seems long enough without it.

As far as we can see, even the following very general statement could be true.

Question 16 Suppose $T_0 \cup T_1 \cup \dots \cup T_s$ is a partition of $[m] \times [n]$, and $a_r, b_r \in \mathbf{N}$ for $r = 1, \dots, s$. Is it true that the ξ_{ij} 's in (3) are NA given

$$\{\xi(T_r) \in [a_r, b_r] \ \forall r \in [s]\}$$

(where $\xi(T) = \sum_{(i,j) \in T} \xi_{ij}$)?

This would be a considerable strengthening of CNA for generalized threshold urn measures.

Let us also just mention one possible approach to Question 13. Recall that for $\mu, \nu \in \mathcal{M}(\{0, 1\}^n)$, μ stochastically dominates ν (written $\mu \succeq \nu$) if $\mu(\mathcal{A}) \geq \nu(\mathcal{A})$ for each increasing $\mathcal{A} \subseteq \{0, 1\}^n$, and that μ has the *normalized matching property* if, with X chosen according to μ and $\xi = |X|$, $\mu(\cdot | \xi = k)$ is stochastically increasing in k (meaning, of course, that $\mu(\cdot | \xi = k) \succeq \mu(\cdot | \xi = l)$ whenever $k > l$). It is not too hard to show (this is somewhat like the derivation of CNA from CNC in [6]) that CNA for generalized interval urn measures would follow from a positive answer to

Question 17 Is it true that for any $\sigma \in [n]^{[m]}$ with law given by (6) and Q as in (7), the law of $\sigma^{-1}(K)$ given Q has the normalized matching property?

Finally we turn to the conjectures of Farr (unpublished circa 2004; see [18]) and Welsh [18] mentioned at the end of Section 1. To put these in our framework, we add an urn Λ and assume

$$\Pr(\sigma(i) = j) = p \quad \forall i \in [m], j \in [n].$$

(So $\Pr(\sigma(i) = \Lambda) = 1 - np$.) Let $\mathcal{I} \subseteq 2^{[m]}$ be decreasing and set $\mathcal{A}_j = \{\sigma^{-1}(j) \in \mathcal{I}\}$ and $\mathcal{A}_J = \cap \{\mathcal{A}_j : j \in J\}$. Then Farr's conjecture (somewhat rephrased) is

Conjecture 18 If G is a graph on $[m]$ and \mathcal{I} is the collection of independent sets of G , then for any disjoint $I, J, K \subseteq [n]$, $\mathcal{A}_I \downarrow \mathcal{A}_J$ given \mathcal{A}_K .

It's not clear why this should require that \mathcal{I} be of the type described, and Welsh's conjecture was that the same conclusion holds for an arbitrary \mathcal{I} . Here we sketch a counterexample to this stronger version. At present we don't see how to extend to a counterexample to Conjecture 18, though we feel that this too is likely to be false.

Example. Let $n = 3$ and $p = 1/3$ (so we don't need Λ). Let $M \cup A \cup B \cup C$ be a partition of $V := [m]$ with $|M| = s$ (large) and $|A| = |B| = |C| = t = s + 3$. Let $\mathcal{I}_1 = 2^{V \setminus M}$,

$$\mathcal{I}_2 = \{X \subseteq V : |X \cap M| < .4|M| \text{ and } X \text{ meets at most two of } A, B, C\},$$

and $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$. Then, we assert,

$$\Pr(\mathcal{A}_{\{3\}}) \Pr(\mathcal{A}_{\{1,2,3\}}) > \Pr(\mathcal{A}_{\{1,3\}}) \Pr(\mathcal{A}_{\{2,3\}}),$$

which contradicts Welsh's conjecture (with $I = \{1\}$, $J = \{2\}$ and $K = \{3\}$). We omit the precise calculations; roughly, with $\alpha = (2/3)^t$ and $c = (3/2)^3$, we have (as $t \rightarrow \infty$)

$$\Pr(\mathcal{A}_L) \sim \begin{cases} (c+3)\alpha & \text{if } |L| = 1 \\ 6\alpha^2 & \text{if } |L| = 2 \\ 6\alpha^3 & \text{if } |L| = 3. \end{cases}$$

Acknowledgment Parts of this work were carried out while the authors were visiting the Isaac Newton Institute and while the first author was visiting MIT. The hospitality of both is gratefully acknowledged.

References

- [1] J. Borcea, P. Brändén and T.M. Liggett, Negative dependence and the geometry of polynomials, *J. Amer. Math. Soc.* **22** (2009), 521-567.
- [2] F. Brenti, Log-concave and unimodal sequences in algebra, combinatorics, and geometry: an update, *Contemporary Math.* **178** (1994), 71-89.
- [3] Y. Choe and D. Wagner, Rayleigh Matroids, *Combin. Probab. Comput.* **15** (2006), 765-781 (arXiv:math.CO/0307096v3).
- [4] P. Doyle and J.L. Snell, *Random Walks and Electric Networks*, Carus Mathematical Monographs **22**, MAA, Washington DC, 1984.
- [5] D. Dubhashi and D. Ranjan: Balls and bins: a study in negative dependence, *Random Structures & Algorithms* **13** (1998), 99-124.
- [6] T. Feder and M. Mihail, Balanced matroids, pp. 26-38 in *Proc. 24th STOC*, ACM, 1992.
- [7] G.H. Hardy, J.E. Littlewood and G. Pólya, *Inequalities*, 2nd edition, Cambridge University Press, Cambridge, 1952.
- [8] O.J. Heilmann and E.H. Lieb, Monomers and dimers, *Phys. Rev. Letters* **24** (1970), 1412-1414.
- [9] O.J. Heilmann and E.H. Lieb, Theory of monomer-dimer systems, *Comm. Math. Physics* **25** (1972), 190-232.
- [10] J. Kahn and M. Neiman, Negative correlation and log-concavity, *Random Structures & Algorithms*, to appear.
- [11] J. Kahn and M. Neiman, A strong log-concavity property for measures on Boolean algebras, submitted.
- [12] H. Kunz, Location of the zeros of the partition function for some classical lattice systems, *Phys. Lett. (A)* (1970), 311-312.
- [13] L. Lovász and M.D. Plummer, *Matching Theory*, North Holland, Amsterdam, 1986.
- [14] M. Neiman, Ph.D. Thesis, Rutgers University, 2009.

- [15] R. Pemantle, Towards a theory of negative dependence, *J. Math. Phys.* **41** (2000), 1371-1390.
- [16] R.P. Stanley, Log-concave and unimodal sequences in algebra, combinatorics, and geometry, pp. 500-535 in *Graph theory and its applications: East and West (Jinan, 1986)*, *Ann. New York Acad. Sci.* **576**,
- [17] D. Wagner, Negatively correlated random variables and Mason's conjecture for independent sets in matroids, *Ann. Comb.* **12** (2008), 211-239.
- [18] D. Welsh, Harris's inequality and its descendents, lecture at Isaac Newton Inst., June 2008, <http://www.newton.ac.uk/programmes/CSM/seminars/062310005.pdf>.